ISRCS 2010

www.inl.gov/isrcs

Recovering from Cyber Attacks using Switching Control and Passivity Indices

Michael J. McCourt and Panos J. Antsaklis

August 11, 2010

Vehicle Dynamics

- The basis for the rendezvous and flocking goals in vehicle coordination is the simple problem of output synchronization
  - Typically local and networked control is used for synchronization
  - The consensus protocol can be used to control the network
- Local control can be used to force each vehicle to be passive (Arcak, 2007)
- A vehicle model that is an Euler-Lagrange system with nonholonomic constraints can be transformed into a passive linear system using local feedback (Yu and Antsaklis, 2010)

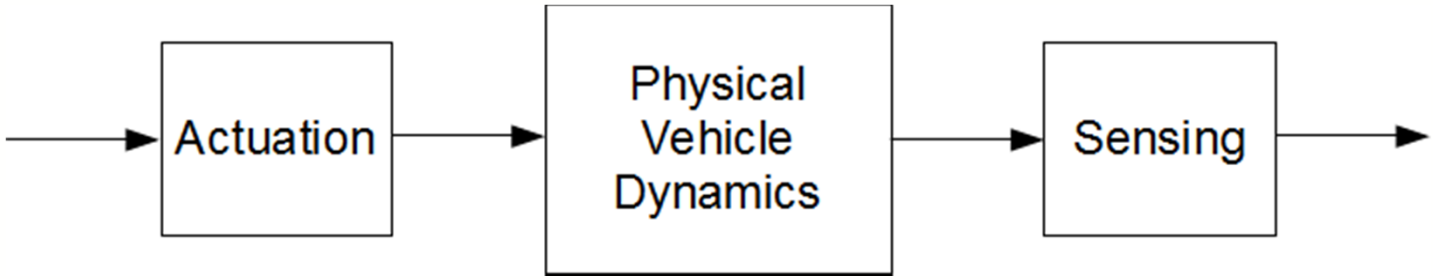
$$\dot{x} = -ax + u$$

$$y = x$$

- However, with cyber attacks, the system may not be passive anymore

Cyber Attacks

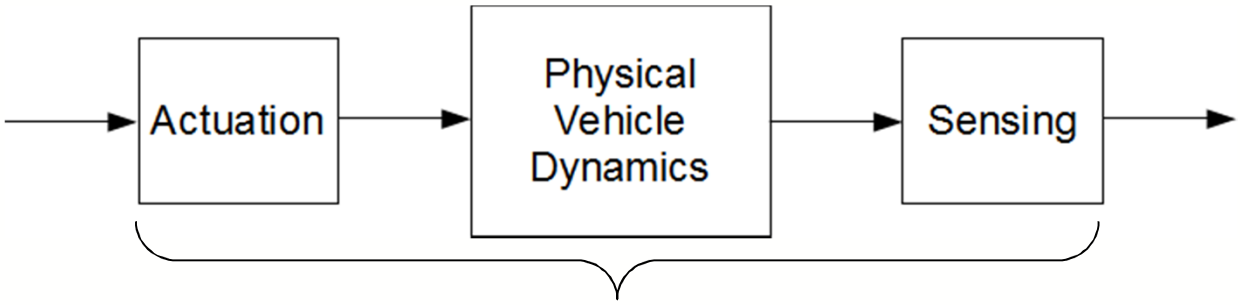
- Cyber attacks can compromise communication between systems or can compromise digital subsystems on a single vehicle
- The approach presented in this talk can be used to compensate for compromised subsystems
  - Consider a vehicle with actuation and sensing
  - The actuation or sensing may become compromised by cyber attacks



Types of Attacks

- Attacks can have varying severity from negligible attacks (such as the addition of noise) to more severe attacks that force entire systems offline
- Typically a single vehicle cannot recover from a severe attack

- This approach can be used to compensate for moderate attacks where the system input and output are still dynamically related



We assume that the compromised system is still a dynamical system but described by a different model that can be nonlinear with unstable or non-minimum phase dynamics

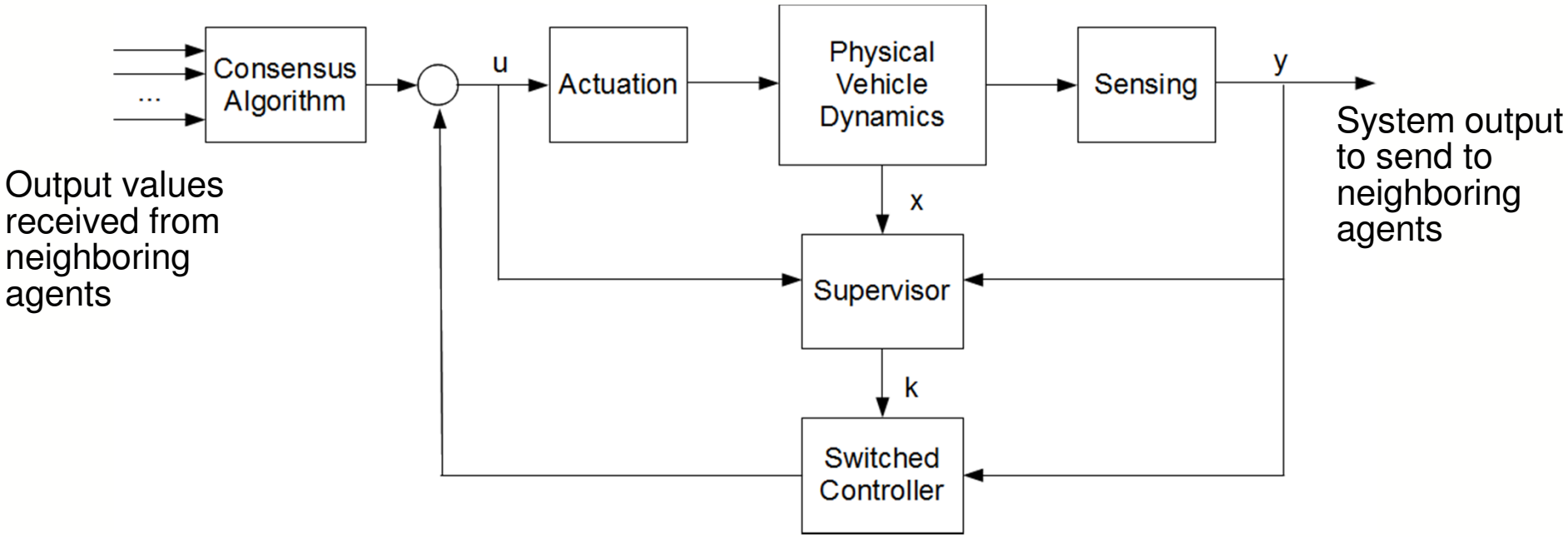
$$\dot{x} = f(x,u)$$
$$y = h(x,u)$$

Supervisory Control Scheme

- The compensation scheme is implemented on an independent wired controller on each vehicle
  - Doesn't use wireless communication so is immune to cyber attacks
  - Contains both a controller (see figure) and a supervisor to force switches in the switched control system
- The controller monitors the possibly compromised input and output of the vehicle and has independent access to the state of the vehicle
  - Uses this data to assess whether the given system is passive and can identify (conservative) estimates of the passivity indices of a system (to be explained)
  - When the given system is not passive, the controller can switch to another subsystem that can compensate for non-passive dynamics
- The controller is a switched system with a finite number of subsystems

$$\dot{x} = f_k(x,u)$$
$$y = h_k(x,u)$$
$$k \in \{1,\ldots,M\}$$

Supervisory Control Scheme

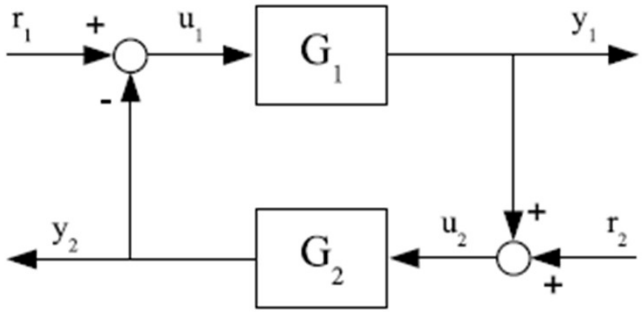


The supervisor tracks the input, output, and state to determine whether the system is passive. If the passivity inequality is violated (see right) the supervisor forces the controller to switch to an appropriate subsystem by changing the discrete index *k*.

$$u^T y < \dot{V}(x)$$

Passivity Index Concept

- Passivity indices can be used to analyze stability of a feedback interconnection
- The feedback of two passive systems is passive and stable



- This analysis can be generalized to systems that aren't passive using passivity indices by assessing the "level" of passivity of a system
- Conceptually, a "nearly" passive system can be compensated by an "excessively" passive system to maintain a stable feedback interconnection

Defining Passivity Indices

- A system is passive if there exists an energy storage function *V(x)* such that the following inequality is satisfied for all inputs *u(t)* and all times *T*

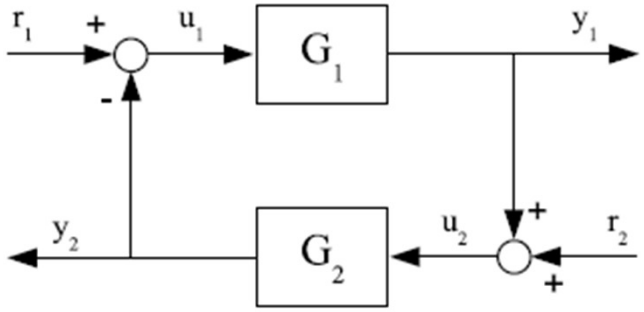
$$\int_0^T u^T y \, dt \geq V\big(x(T)\big) - V(x(0))$$

- A system has passivity indices *ρ* and *v* if there exists an energy storage function *V(x)* such that the following inequality is satisfied for all inputs *u(t)* and all times *T*

$$\int_0^T [(1+\rho v)u^T y - v u^T u - \rho y^T y] \, dt \geq V\big(x(T)\big) - V(x(0))$$

Stability Using Passivity Indices

- Consider the feedback of two dynamical systems G_1 and G_2
  - G_1 has indices ρ_1 and ν_1
  - G_2 has indices ρ_2 and ν_2



- The interconnection is stable if the following matrix is positive definite

$$\begin{bmatrix} (\rho_1 + \nu_2)I & {}^1/_2\,(\rho_2\nu_2 - \rho_1\nu_1)I \\ {}^1/_2\,(\rho_2\nu_2 - \rho_1\nu_1)I & (\rho_2 + \nu_1)I \end{bmatrix} > 0$$

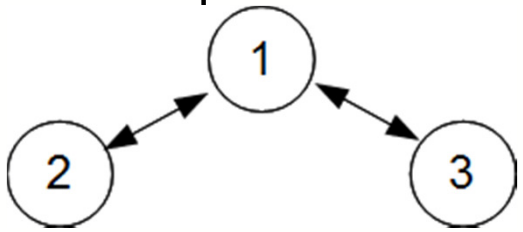
- The concept has been recently applied to switched systems (McCourt and Antsaklis, 2010)

Designing Resilient Switching Controllers

- Several feedback controllers can be designed using passivity indices to compensate for varying levels of malicious attacks
- In many systems there are expected limits to the undesirable dynamics based on physical limitations of the given system
  - For example, the magnitude of unstable modes is bounded by physical limitations
  - Although the magnitudes of the attacks are not known the upper bound is known approximately
- At least one controller should be designed to maintain stability when the dynamics are near the upper bound
- A supervisor can initiate switches in the controller to maintain stability

Example

- Consider a coordinated vehicle problem with the following communication setup



- The nominal dynamics of each subsystem are linearized and made passive (a>0) by local control

$$\dot{x}_i = -ax_i + u_i$$

$$y_i = x_i$$

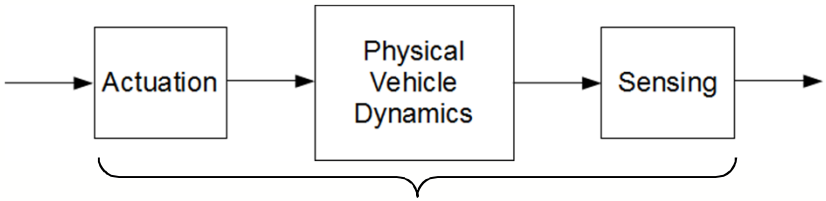
- The consensus protocol is applied for the nominal system to reach output synchronization (*j* is in the set of agents that can communicate with agent *i*)

$$u_i = \sum_{j=1}^{n} (y_i - y_j)$$

- At some point, one of the vehicles (vehicle 2) is subject to a cyber attack that makes its dynamics unstable

Cyber Attacks

- When a cyber attack compromises the actuation or the sensing, the system switches to an unstable system



$$\dot{x}_i = ax_i + u_i$$
$$y_i = x_i$$

- Although the model is the same, the mode is now unstable (*a>0*)
- Being an unknown attack, the value of the mode is unknown but it is known to be bounded above by $\bar{a}$
- A controller can be designed using passivity indices to have indices $v>\bar{a}$ and $\rho>0$ in order to maintain stability
- The attacks happen at unknown times, this was simulated using a uniform random variable over a range of values

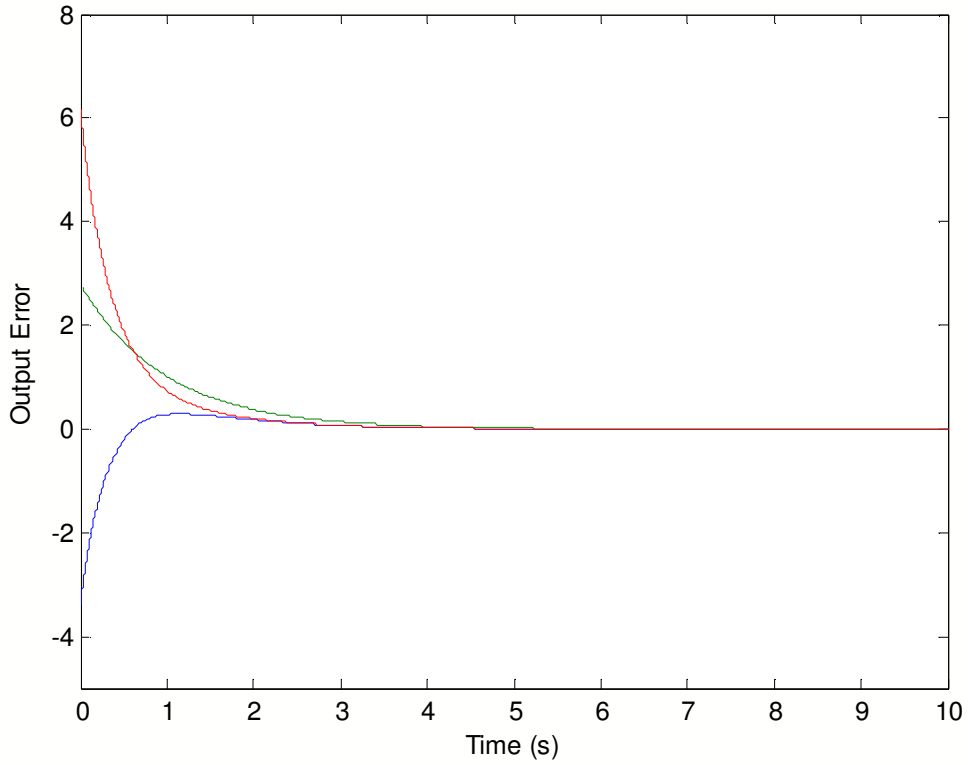
Example - Real Time Recovery

1. The system is stable until a cyber attack occurs
2. The supervisor tracks the passive inequality as well as the inequalities for various sets of passivity indices in order to test passivity or to estimate (conservatively) the current indices
3. When the passive inequality is about the be violated, the controller switches to another subsystem that maintains stability
4. The overall system continues to run stably, typically with reduced performance

- This system was simulated with two ranges of possible attacks and two controllers to compensate for these attacks
- The following graphs show the output error between nodes to show synchronization

Nominal Response

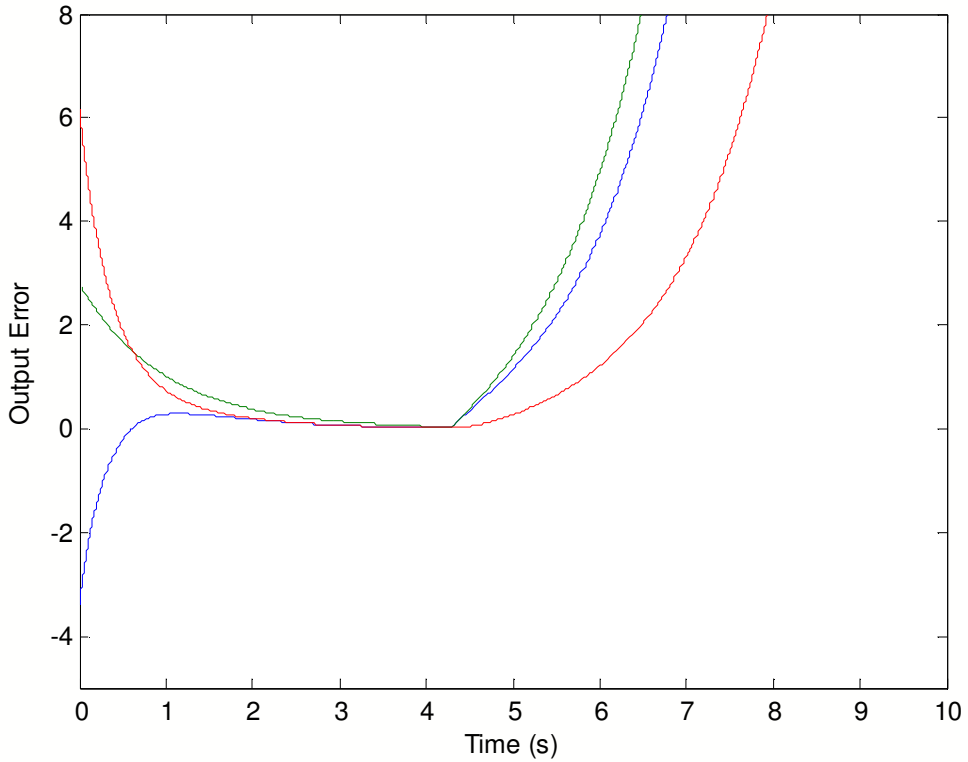
- Without a cyber attack the system behaves as follows where synchronization is achieved rather quickly



- The nominal behavior is for the system to settle in approximately 4.9s

Response with an attack

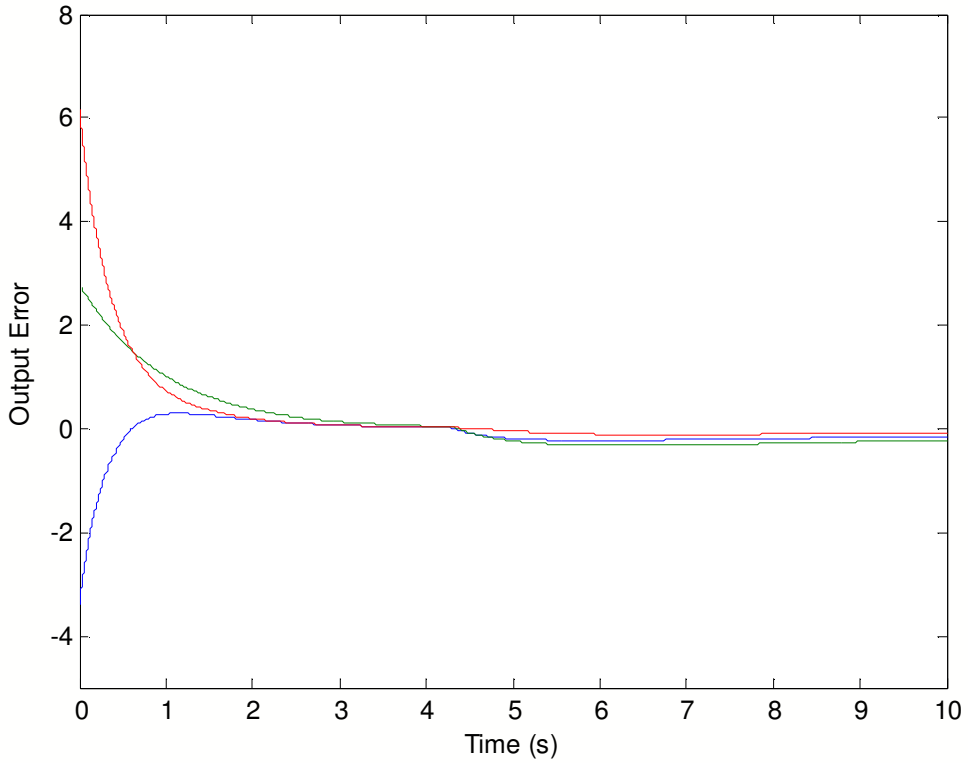
- The following shows the result of a cyber attack when no supervisory control scheme is present



- The single agent has unstable dynamics that skews the other agents towards infinity

Resilient Control

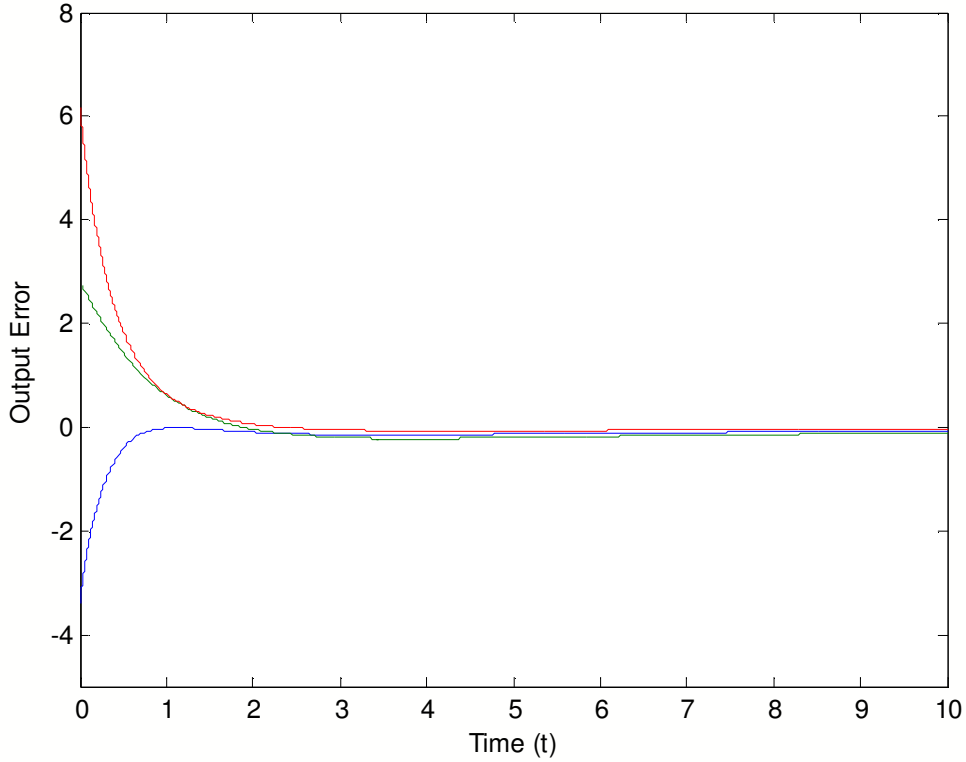
- The following shows the system response when the compensating supervisor is implemented



- The attack happens at 4.3s. The system recovers with reduced performance (settles in 23.3s instead of 4.9s)

Series of Attacks

- In more extreme cases, a serious of attacks of varying strength may occur. In this example, 100 attacks occurred randomly over the time period from 0s to 8s and the system recovered from each attack



Summary

- Passivity indices can be used to asses the level of passivity of a system
- A supervisory control scheme can recover stability by monitoring passivity indices and switching to an appropriate controller
- A coordinated vehicle system can still be synchronized even when the vehicles have been compromised by cyber attacks
  - The passivity index framework can be used for either unstable or non-minimum phase dynamics but not both